Supporting Information

Role of Molecular Interactions and Protein Rearrangement in the Dissociation Kinetics of p38α MAP Kinase Type-I/II/III Inhibitors

Wanli You and Chia-en A. Chang*

Department of Chemistry, University of California at Riverside, Riverside, CA 92521, USA

* Corresponding author: Chia-en A. Chang

Email: chiaenc@ucr.edu

Telephone: (951) 827-7263

Table S1. Simulations performed for each system.

Free DFG-in protein	1. Three CMD runs with different random number seeds (100ns CMD1, CMD2 and CMD3). 2. CMD1, CMD2 and CMD3 were used for compute average RMSF and correlation values.
Free DFG-out protein	1. Three CMD runs with different random number seeds (100ns CMD1, CMD2 and CMD3). 2. CMD1, CMD2 and CMD3 were used for compute average RMSF and correlation values.
SB2 (DFG-in) complex	1. Three CMD runs with different random number seeds (650ns CMD1, 100ns CMD2 and CMD3). 2. One low boost 100ns AMD continued from the first 100ns of CMD1. 3. PSIM search with trajectories for ATP pathway. 4. Two successful 100ns high boost AMD continued from the first 100ns of CMD1. 5. Two 10ns CMD trajectories starting from selected frames of high boost AMD were used to rebuilt the whole smooth dissociation path. 6. CMD2, CMD3 and the first 100ns CMD1 were used for compute average RMSF and correlation values.
SB2 (DFG-out) complex	1. Three CMD runs with different random number seeds (100ns CMD1, CMD2 and CMD3). 2. PSIM search with trajectories for ATP pathway. 3. Two successful 100ns high boost AMD continued from CMD1. 4. Two 10ns CMD trajectories starting from selected frames of high boost AMD were used to rebuilt the whole smooth dissociation path. 5. CMD1, CMD2 and CMD3 were used for compute average RMSF and correlation values.
SK8 complex	1. Three CMD runs with different random number seeds (100ns CMD1, CMD2 and CMD3). 2. PSIM search with trajectories for ATP pathway. 3. Two successful 100ns high boost AMD continued from CMD1. 4. Three 10ns CMD trajectories starting from selected frames of high boost AMD were used to rebuilt the whole smooth dissociation path. 5. CMD1, CMD2 and CMD3 were used for compute average RMSF and correlation values.
BIRB796 complex	1. Three CMD runs with different random number seeds (650ns CMD1, 100ns CMD2 and CMD3).

	2. PSIM search with four trajectories for allosteric pathway and one for ATP pathway. 3. Two 10ns CMD trajectories starting from selected frames of PSIM were used to rebuilt the two whole smooth dissociation paths each. 4. CMD2, CMD3 and the first 100ns CMD1 were used for compute average RMSF and correlation values.
LIG4 complex	1. Three CMD runs with different random number seeds (100ns CMD1, CMD2 and CMD3). 2. PSIM search with trajectories for allosteric pathway and ATP pathway. 3. Three 10ns CMD trajectories starting from selected frames of PSIM were used to rebuilt the whole smooth allosteric path, and two CMD for ATP path. 4. CMD1, CMD2 and CMD3 were used for compute average RMSF and correlation values.

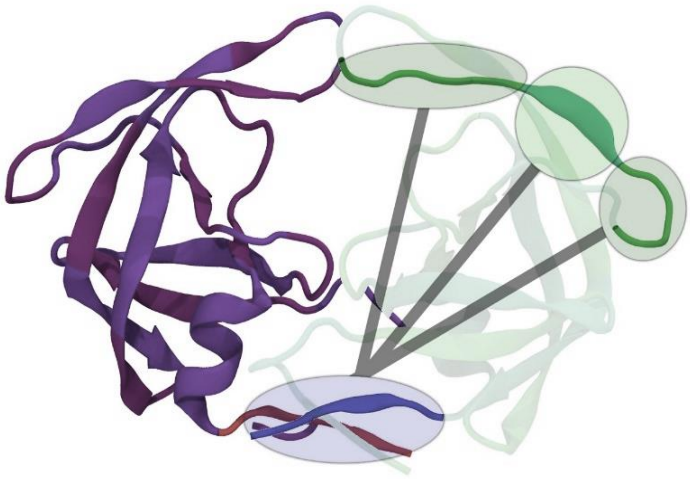


Figure S1. An example of fragmentation of HIV-1 protease (HIVp) for multi-layer internal coordinate definition. Each monomer is considered one group and divided into 21 fragments. The root fragment on monomer 1 is in dark blue shown in the bottom of the protein. The other fragments on monomer 1 are in green shown in the right, and these fragments are defined as the root fragments. Monomer 2 is shown in the left chain of HIVp and its root fragment is in dark red. The root fragments of the two monomers are connected to each other. All fragments including the two root fragments are connected through six pseudo bond, angle and dihedral degrees of freedom. For illustration purposes, the root fragment and three other fragments on monomer 1 are circled. The connections between them are represented by thick gray lines.

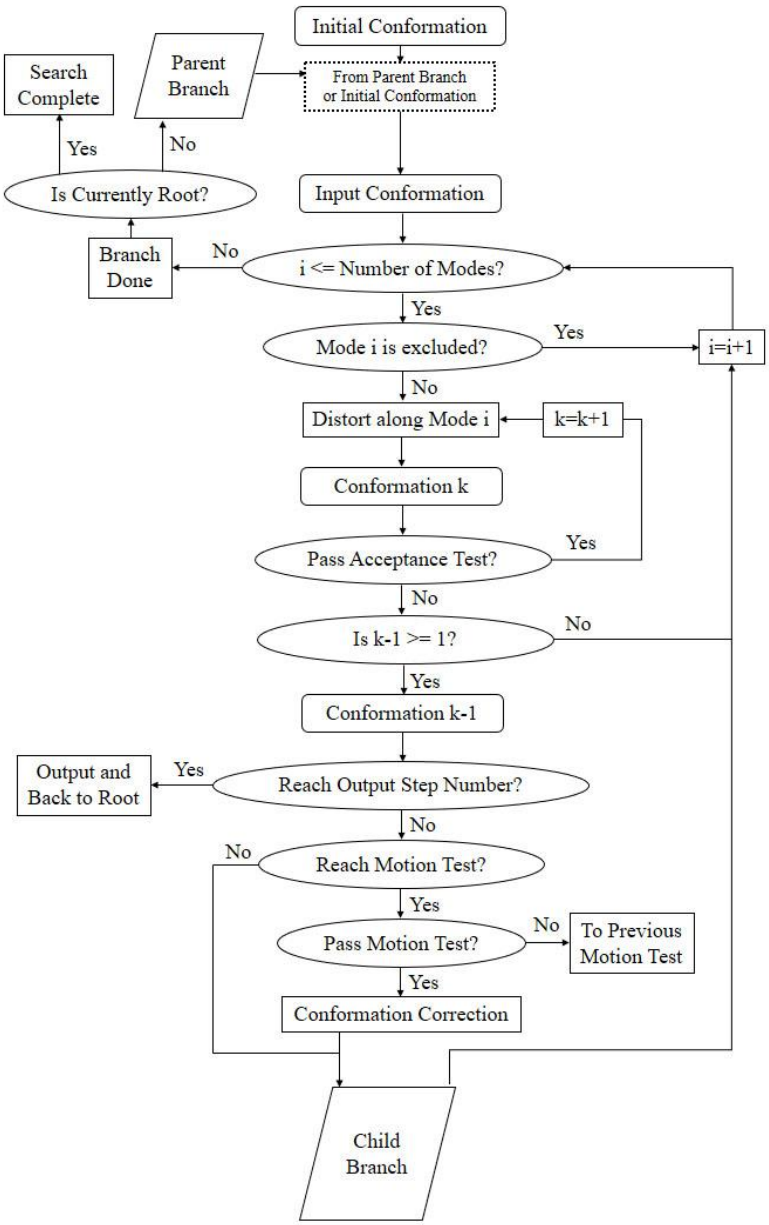


Figure S2. Overview of the algorithm in one search branch used in PSIM. Rounded rectangles represent the conformations, which are new conformations generated by distortions or the input from the parent branch. Rectangles represent the operation modules. Ellipses represent a decision or a test module. PSIM enters a search branch represented by the flowchart from an initial conformation or a conformation produced by its previous search. It starts the loop over principal component (PC) modes presented as i here and postpones the PC mode loop when a direction of a PC mode is done. After certain distortions, PSIM will perform different test to decide if a search should be continued or stopped. A systematic search run is accomplished when all modes/branches have been visited.

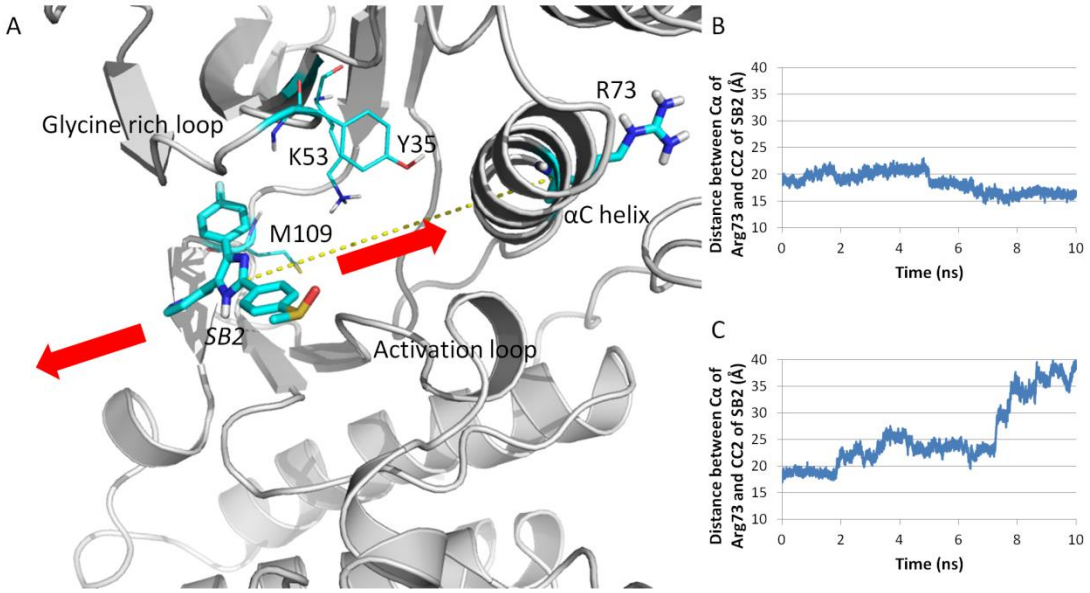


Figure S3. Reconstruction of dissociation path from AMD. Dissociation path of SB2 from p38α with a DFG-in conformation is rebuilt from two 10 ns CMD. (A) SB2 in one of the two 10 ns CMD moves towards inside the cavity, while SB2 in the other CMD moves towards outside. Arg73 and SB2 are shown in bold licorice structure, key interacting residues are shown in thin licorice structure. (B) SB2 moving inside the cavity indicated by the decreasing distance between SB2 and Arg73. (C) SB2 moving outside the cavity indicated by the increasing distance between SB2 and Arg73.

Section 1

To test the suitability and advantage of AMD, we first compared the results of CMD, low boost AMD and high boost AMD for SB2 (DFG-in). The average RMSD of heavy backbone atoms in trajectories relative to crystal structure are 2.05 for 100ns CMD, 3.05 for 100ns low boost AMD and 3.26 for high boost AMD, respectively. Figure S4 shows protein in low and high boost AMD didn't move too far from the crystal structure. Although in high boost AMD, the deviation increases significantly after SB2 dissociated around 64ns, indicating the role of ligand in stabilizing protein structure in AMD.

Figure S5 shows the projection of three trajectories (100ns CMD, 100ns low boost AMD and 100ns high boost AMD) on the first and second principal component vectors (PC1 and PC2) built from the Cα atoms of p38α in high boost AMD trajectory. In the CMD trajectory, PC1 and PC2 describe 23% and 12%, respectively, of the total variance of the motions in the simulation. It is clear from Figure S5A that CMD is trapped in the basin of crystal structure and never gets out. Figure S5B shows that the low boost AMD simulation does not explore the amount of conformational space that

the high boost AMD simulation does and remains trapped around the crystal structure. The high boost AMD simulation (Figure S5C) exhibits a rather broad pathway from the crystallographic basin (-35, −20) to the region (15, 35) which represents SB2 outside binding cavity and is not present in the 100ns CMD or low boost AMD simulations (Figure S5A,B). In sum, high boost AMD can successfully simulate the dissociation of type-I ligand and maintain an overall stable structure through simulation.

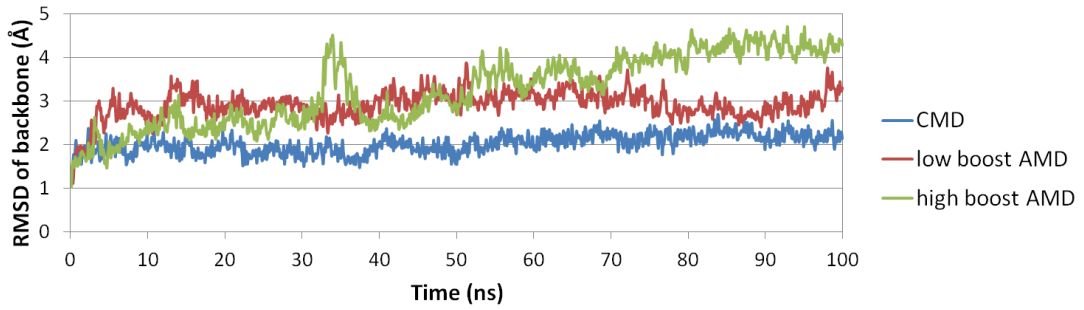


Figure S4. RMSD of heavy backbone atoms in trajectories of CMD, low boost AMD and high boost AMD relative to crystal structure.

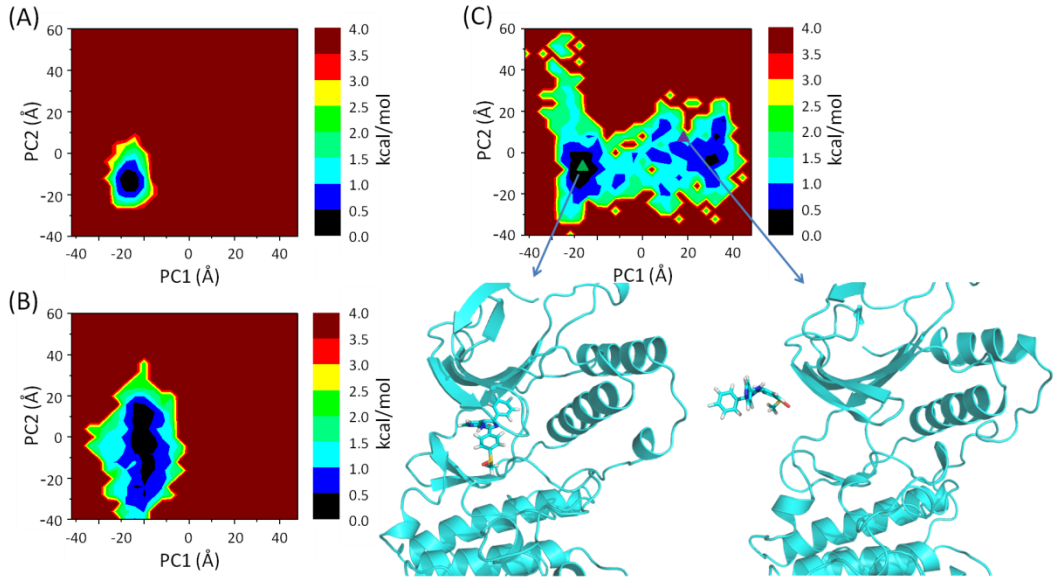


Figure S5. The free energy principal component projection of (A) 100ns CMD, (B) 100 ns low boost AMD, and (C) 100 ns high boost AMD onto (PC1, PC2) defined by the 100 ns high boost AMD. CMD and low boost AMD are trapped in the basin of crystal structure, while high boost AMD explores much larger conformational space.

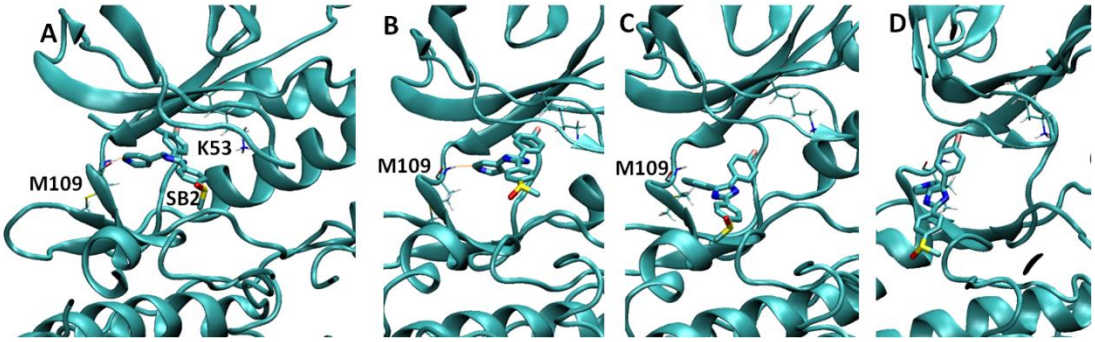


Figure S6. Snapshots from PSIM trajectory of dissociation process of SB2 (DFG-in). SB2 is shown in bold licorice structure. Hydrogen bonds between SB2 and p38α are shown with dashed lines. Key interacting residues are shown in thin licorice structure which are also found in Figures 3ABCD in the main text. First the hydrogen bond between SB2 and Lys53 breaks (Fig S6A; same as Figure 3A. Ligand RMSD: 2.3Å), followed by the motion of the 4-methylsulfinylphenyl group (Fig S6B; same as Figure 3B. Ligand RMSD: 1.8 Å). Then the second hydrogen bond between pyridine nitrogen and Met109 breaks (Fig S6C; same as Figure 3C. Ligand RMSD: 1.2 Å) and finally, the ligand is outside the edge of the binding cavity (Fig S6D; Figure 3D. Ligand RMSD: 1.9 Å) and eventually diffuses away. Note that trajectories from PSIM and AMD were first aligned using the secondary structures of p38 before computing ligand RMSD.

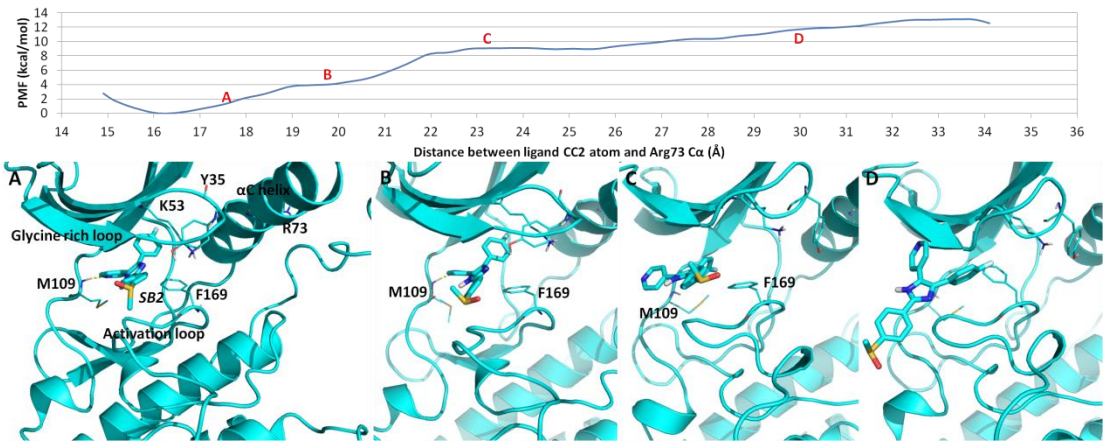


Figure S7. PMF of dissociation process of SB2 (DFG-out) and the selected snapshots from US. SB2 is shown in bold licorice structure, key interacting residues are shown in thin licorice structure, hydrogen bonds between SB2 and p38α are shown in dash line. (A) SB2 breaks hydrogen bond with Lys53 side-chain and stacking interaction with rotation of Tyr35. (B) 4-methylsulfinylphenyl group of SB2 diffuses towards outside the cavity, fluorophenyl ring of SB2 moves out of the hydrophobic pocket and forms stacking interaction with Phe169. (C) SB2 breaks hydrogen bond with Met109 stacking interaction with Phe169. (D) SB2 is outside the binding cavity.

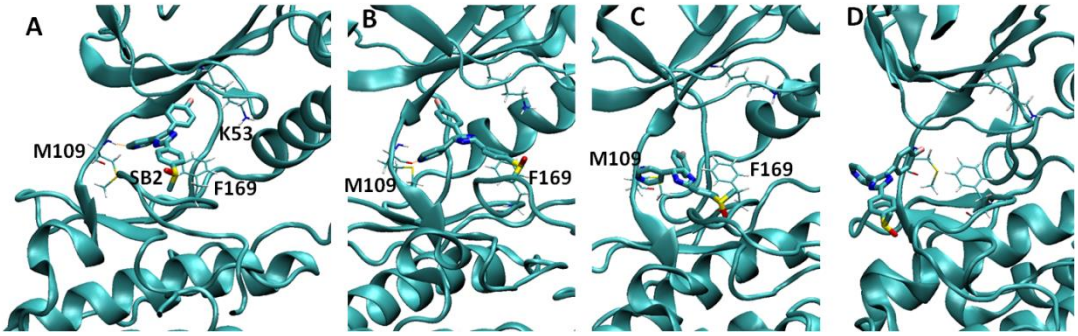


Figure S8. Snapshots from PSIM trajectory of dissociation process of SB2 (DFG-out). SB2 is shown in bold licorice structure. Hydrogen bonds between SB2 and p38α are shown with dashed lines. Key interacting residues are shown in thin licorice structure which are also found in Figures S7ABCD in the SI. First the hydrogen bond between SB2 and Lys53 breaks (Fig S8A; same as Figure S7A. Ligand RMSD: 3.0 Å), followed by 4-methylsulfinylphenyl group of SB2 diffuses towards outside the cavity, fluorophenyl ring of SB2 moves out of the hydrophobic pocket and forms stacking interaction with Phe169. (Fig S8B; same as Figure S7B. Ligand RMSD: 3.5 Å). SB2 breaks hydrogen bond with Met109 and breaks stacking interaction with Phe169. (Fig S8C; same as Figure S7C. Ligand RMSD: 2.7 Å) and finally, the ligand is outside the binding cavity (Fig S8D; Figure S7D. Ligand RMSD: 3.2 Å). Note that trajectories from PSIM and AMD were first aligned using the secondary structures of p38 before computing ligand RMSD.

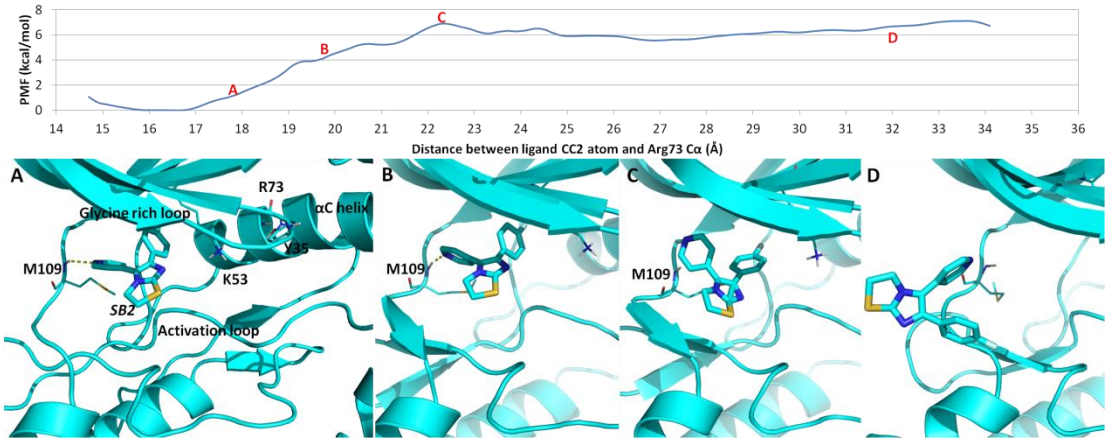


Figure S9. PMF of dissociation process of SK8 and selected snapshots from US. SK8 is shown in bold licorice structure, key interacting residues are shown in thin licorice structure, hydrogen bonds between SK8 and p38α are shown in dash line. (A) SK8 breaks hydrogen bond with Lys53 side-chain and stacking interaction with rotation of Tyr35. (B) Fluorophenyl ring of SK8 moves out of the hydrophobic pocket. (C) SK8 breaks hydrogen bond with Met109. (D) SB2 is outside the binding cavity.

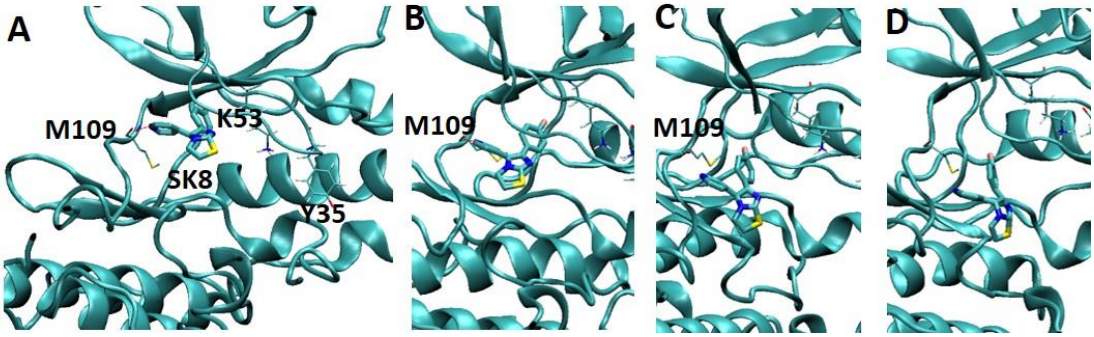


Figure S10. Snapshots from PSIM trajectory of dissociation process of SK8. SK8 is shown in bold licorice structure. Hydrogen bonds between SK8 and p38α are shown with dashed lines. Key interacting residues are shown in thin licorice structure which are also found in Figures S9ABCD in the SI. First the hydrogen bond between SK8 and Lys53 breaks and Tyr35 moves away (Fig S10A; same as Figure S9A. Ligand RMSD: 1.6 Å), followed fluorophenyl ring of SK8 moving out of the hydrophobic pocket (Fig S10B; same as Figure S9B. Ligand RMSD: 2.3 Å). Then SK8 breaks hydrogen bond with Met109 (Fig S10C; same as Figure S9C. Ligand RMSD: 3.4Å) and finally, the SK8 is outside the binding cavity (Fig S10D; same as Figure S9D. Ligand RMSD: 3.5 Å). Note that trajectories from PSIM and AMD were first aligned using the secondary structures of p38 before computing ligand RMSD.

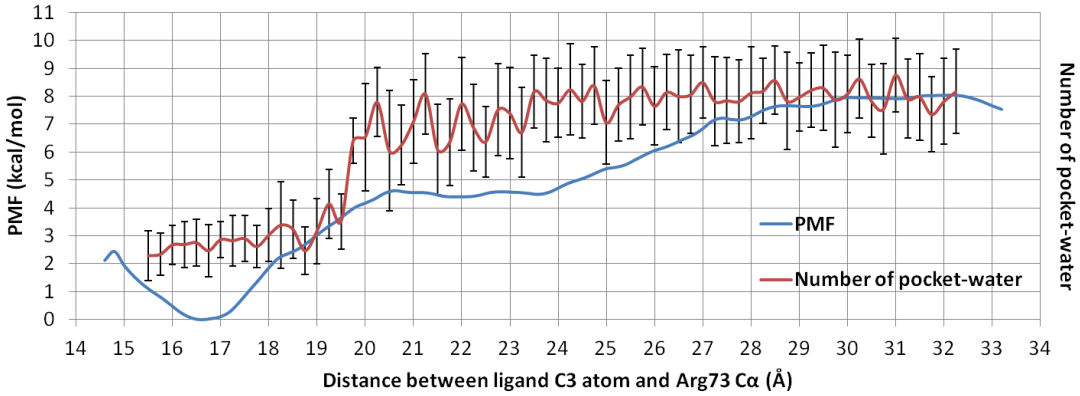


Figure S11. PMF of SB2 (DFG-in) dissociation (blue curve) and profile of number of pocket-water (red curve) as a function of the RC distance. Standard deviation of number of pocket-water is used as error bar. To define the pocket, we first selected the residues that compose the pocket and also identified the center of the pocket. All water molecules within an effective cut off distance 5Å of the center of pocket and the residues inside the pocket were counted as “pocket-water molecules".

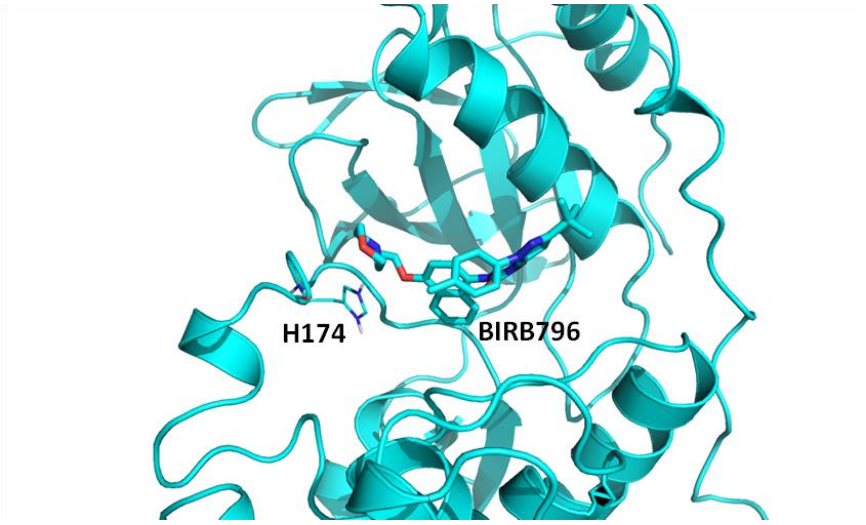


Figure S12. Snapshots from PSIM trajectory of dissociation process of BIRB796. BIRB796 and His174 are shown in bold and thin licorice structures, respectively. Similar to Figure 5B, after BIRB796 breaks the two hydrogen bonds between Glu71 and Asp168, the ligand moves outwards and then forms stacking interaction with His174 and moves away.

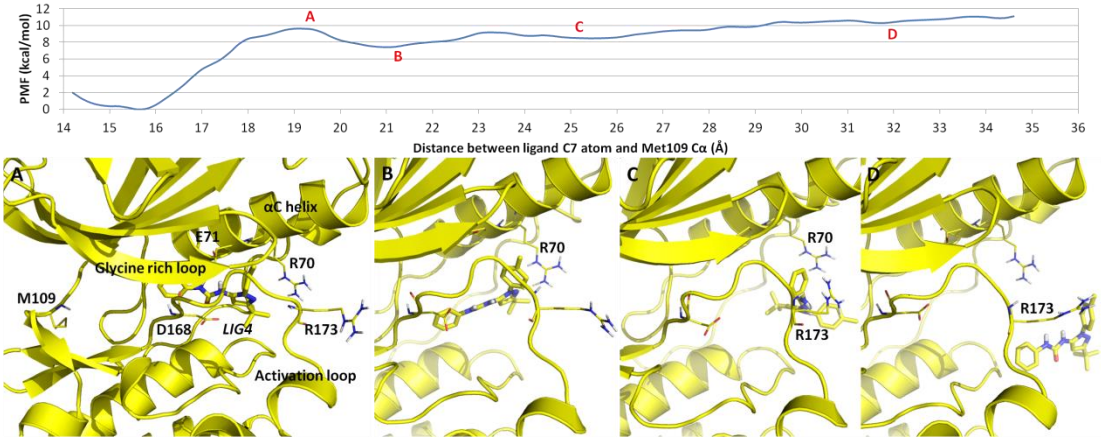


Figure S13. PMF of dissociation process of LIG4 along allosteric pathway and selected snapshots from US. LIG4 is shown in bold licorice structure, key interacting residues are shown in thin licorice structure, hydrogen bonds between LIG4 and p38α are shown in dash line. (A) Cleft opens up, breaking hydrogen bonds between urea group of LIG4 and Glu71 and Asp168. (B) LIG4 diffuses out, forms hydrogen bond with Arg70. (C) LIG4 breaks hydrogen bond with Arg70, forms ring-ion stacking interaction with Arg173. (D) LIG4 breaks stacking interaction with Arg173 and diffuses away.

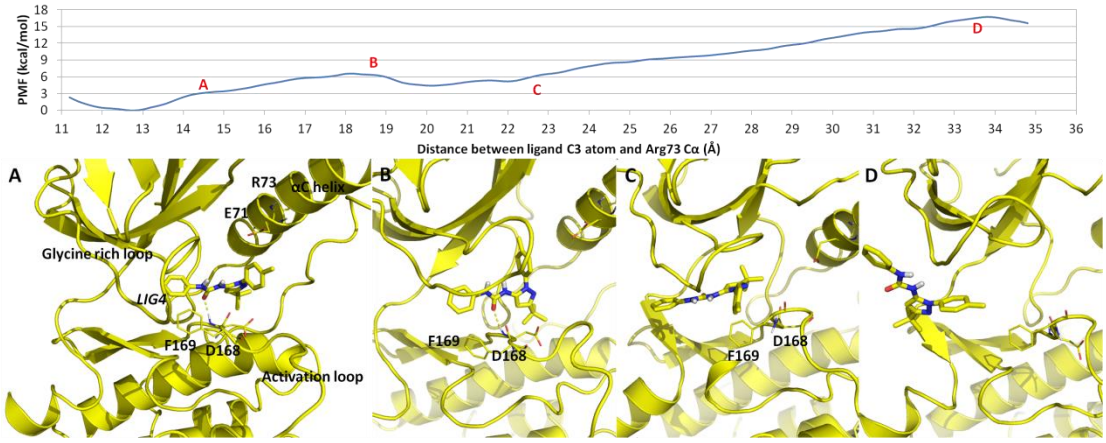


Figure S14. PMF of dissociation process of LIG4 along ATP pathway and selected snapshots from US. LIG4 is shown in bold licorice structure, key interacting residues are shown in thin licorice structure, hydrogen bonds between LIG4 and p38α are shown in dash line. (A) Cleft opens up, breaking hydrogen bonds between urea group of LIG4 and Glu71. LIG4 forms stacking interaction between phenyl group and Phe169. (B) LIG4 enters ATP binding site. (C) LIG4 breaks hydrogen bond with Asp168, Phe169 switches to form stacking interaction with toluene moiety. (D) LIG4 breaks stacking interaction with Phe169 and diffuses away.

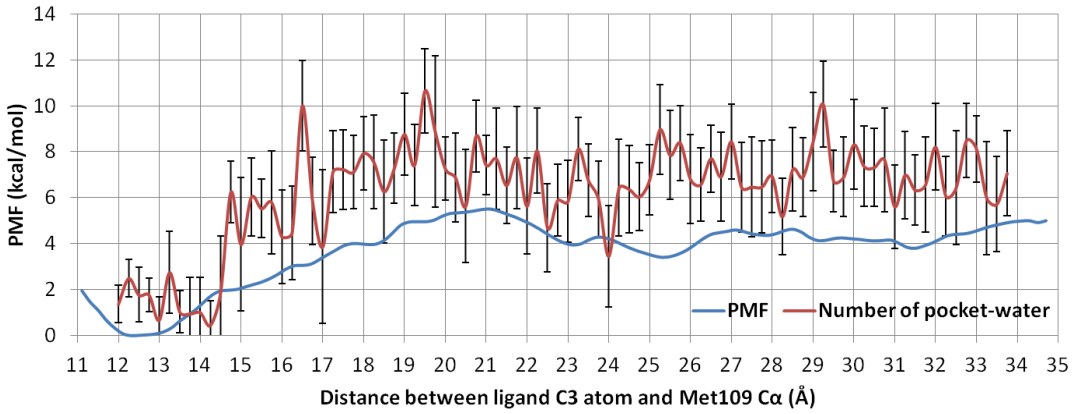


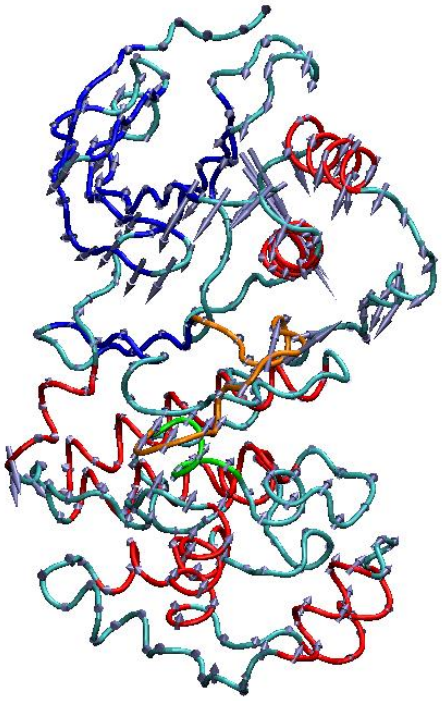
Figure S15. PMF of BIRB796 dissociation (blue curve) and profile of number of pocket-water (red curve) as a function of the RC distance. Standard deviation of number of pocket-water is used as error bar. To define the pocket, we first selected the residues that compose the pocket and also identified the center of the pocket. All water molecules within an effective cut off distance 5Å of the center of pocket and the residues inside the pocket were counted as “pocket-water molecules".

Section 2

It turns out that in free DFG-out protein, the most principal motion is the hinge motion between N and C lobes is mainly between activation loop and αC helix, which are on the sides of allosteric path. In free DFG-in protein, the hinge motion is mainly between glycine-rich loop (β1, L4, β2) and L9/αD helix, which are on the way of ATP path. Both PCA results suggest that hinge movement may be important in ligand dissociation.

Comparing to free protein, p38α complexes retain strong correlation in hinge region, while the correlation between ligand and protein are mostly not obvious due to the limited power of conformational sampling using CMD (Figure S19). Even though ligands stays steadily in energy basin of crystal bound position, there is still correlation between SB2 (both DFG-in and DFG-out) and the hinge and activation loop (P loop and P+1 loop). SB2 (DFG-out) has stronger correlation with hinge region and activation loop, especially with the activation loop because of the extra stacking interaction between the 4-methylsulfinylphenyl group of SB2 and Phe169, which is also reflected in large fluctuation in activation loop during CMD simulation. BIRB796 and LIG4 don't show any strong correlation between ligand and protein due to the grip of two hydrogen bonds from N and C lobes. Interestingly, in BIRB796 complex, there is strong correlation between L16 loop/α2L14 helix region and other part of protein, which confirms the concerted motion between L16 loop/α2L14 helix region and hinge region, as seen in PCA motion in Figure S16. It is possible that L16 loop/α2L14 helix region plays a role in helping the hinge region move.

free DFG-in



free DFG-out

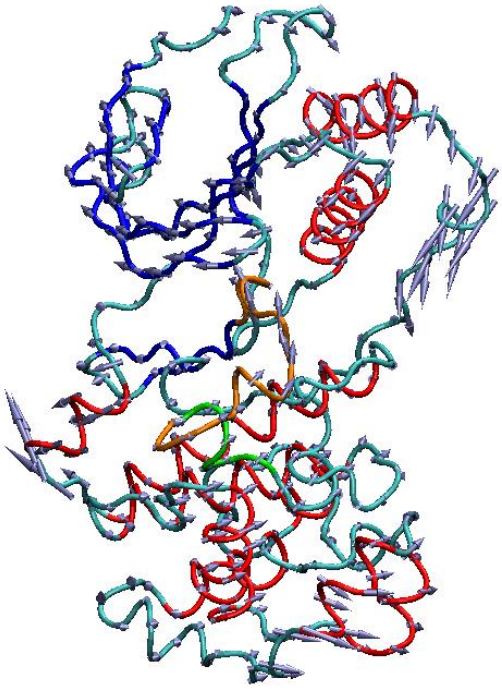


Figure S16. The first PC modes of free DFG-in and DFG-out proteins from CMD simulations. The gray arrows indicate the local direction and magnitude of movement. α-helix is colored in red, β-sheet is colored in blue, loop is colored in cyan, activation loop is colored in orange, P+1 substrate site is colored in green.

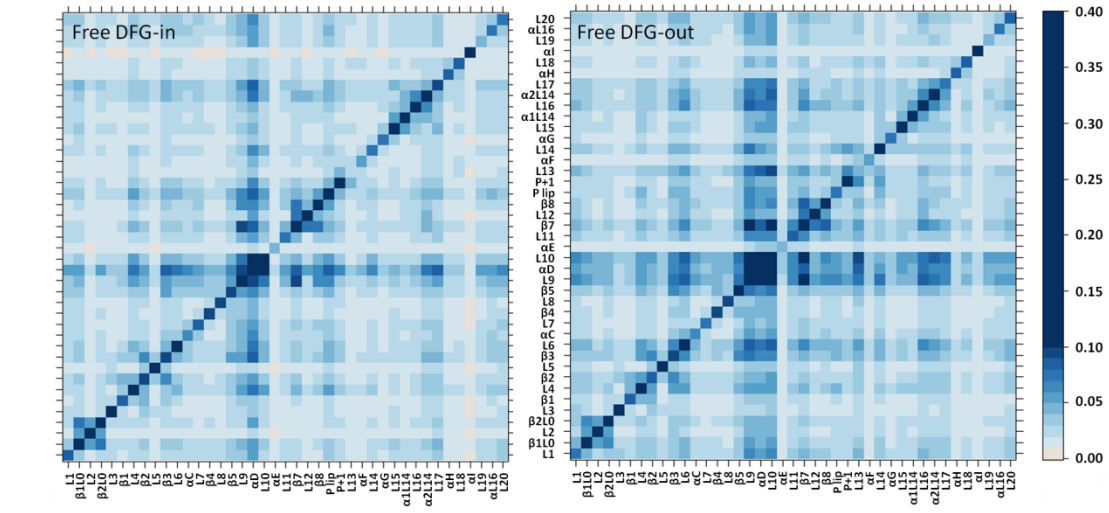


Figure S17. Correlation maps of free DFG-in and DFG-out proteins from CMD simulations.

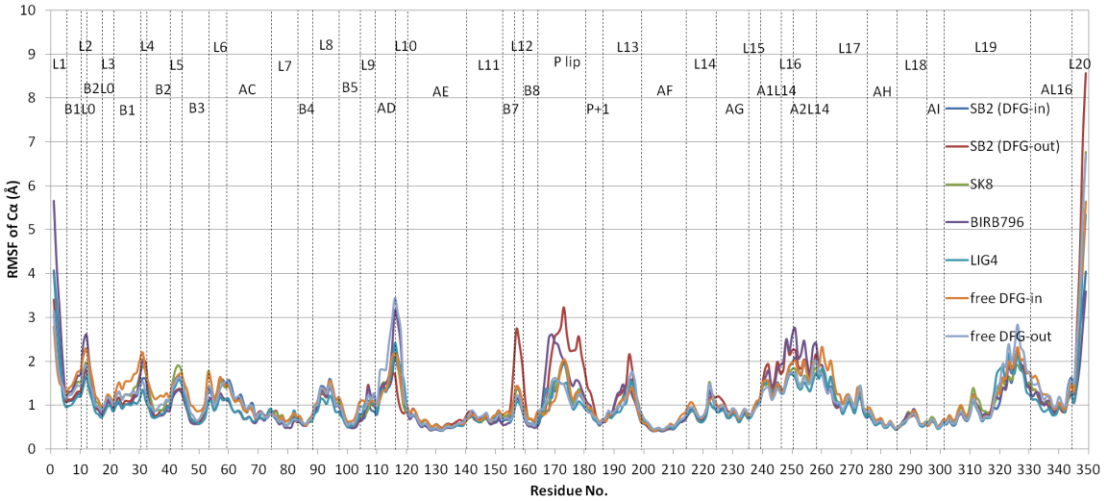
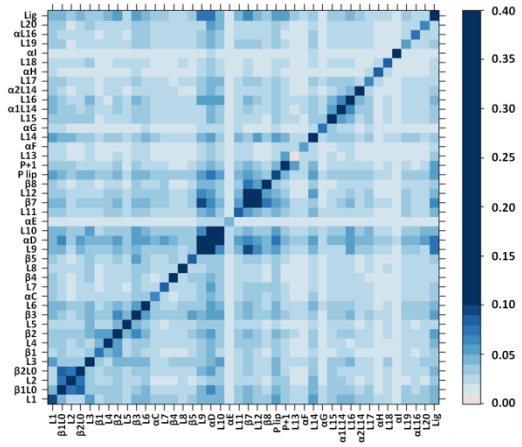
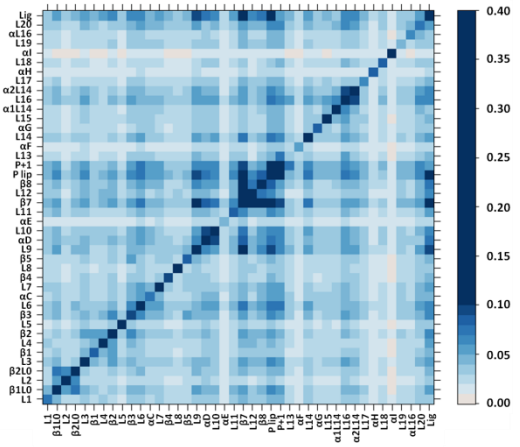


Figure S18. RMSF of Cα of p38α complexes and free proteins from CMD simulations. Residues are divided into different different fragments basing on the secondary structure of protein. For example, L1 indicated Loop 1, B1L0 indicates β-sheet 1L0, AC indicates α-helix C, P lip indicates activation loop, P+1 indicated P+1 substrate site.

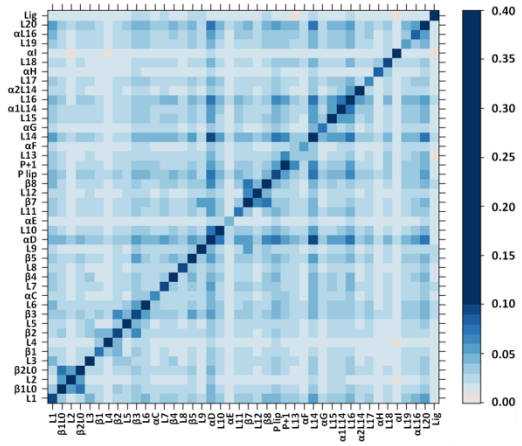
SB2 (DFG-in)



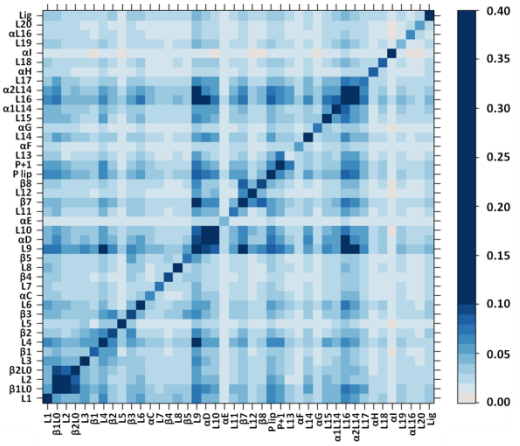
SB2 (DFG-out)



SK8



BIRB796



LIG4

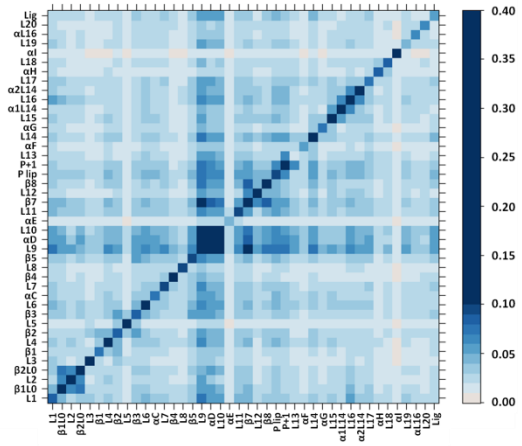


Figure S19. Correlation maps of p38α complexes from CMD simulations.

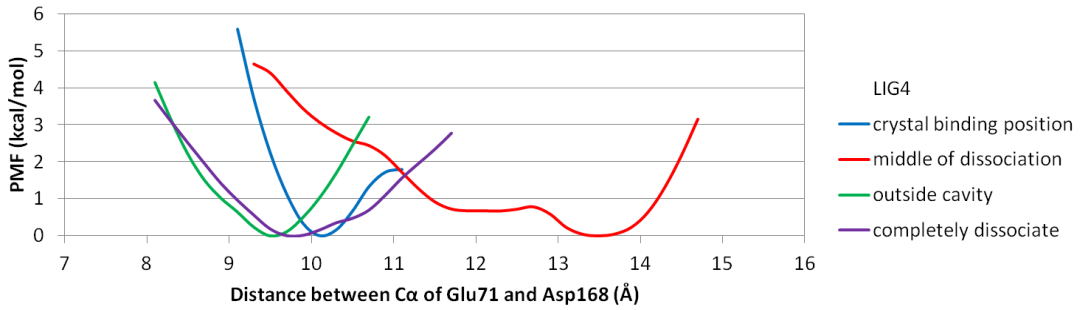


Figure S20. Free energy change along hinge movement of p38α in DFG-out conformation at different stages of dissociation of LIG4. Distance between Cα of Glu71 and Asp168 is used as RC.

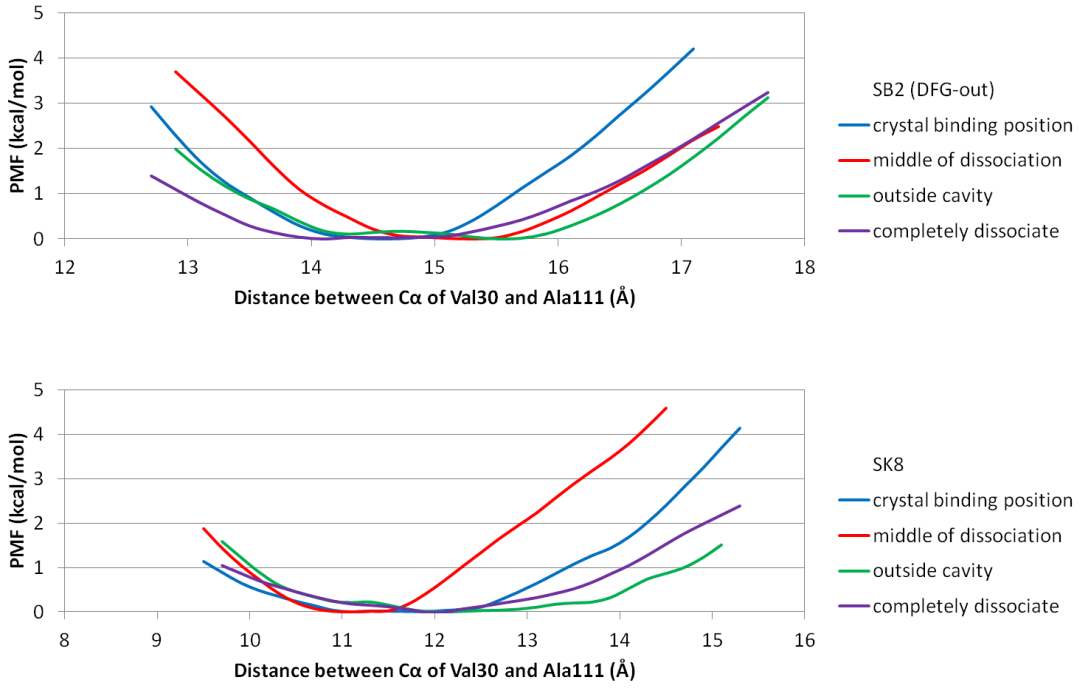


Figure S21. Free energy change along hinge movement of p38α at different stages of dissociation of SB2 (DFG-out) and SK8. Distance between Cα of Val30 and Ala111 is used as RC.